# Internet-Enabled Geotechnical Data Exchange

David E. Kosnik[1]

[1]Research Engineer, Northwestern University Infrastructure Technology Institute, 2133 Sheridan Road, Evanston, Illinois 60208; dkosnik@northwestern.edu

**ABSTRACT:** Surveillance of large geotechnical projects requires autonomous collection of data from a wide range of instrument types. An equally large range of data formats are employed by these measurement systems. A data exchange protocol is needed to make field data available for interpretation on a variety of platforms. This paper presents the results of the development of a robust Internet-enabled framework for the exchange of engineering data. Internet-enabled data exchange technology provides a standard for engineering data acquired from any combination of instruments by autonomously parsing data into a searchable database for archiving and distribution. A set of strategies and software have been deployed to create layers of abstraction that allow data files from different commercial and in-house data acquisition packages to meet standards for timekeeping, archiving, and data display. The data are stored in an Internet-accessible database and distributed to stakeholders via a password-protected Web site. The data exchange abstractions in this system allow it to be employed with an extremely wide variety of sensing techniques and data types, and the storage of data in a relational database rather than flat files provides for improved scalability in the types and volumes of data to be interpreted. These technologies have been employed in the continuous remote monitoring of a number of facilities over the past seven years.

## INTRODUCTION

A wide variety of sensing devices and techniques are currently employed in geomechanical measurements today, including tiltmeters, inclinometers, automated survey stations, piezometers, displacement sensors, and time-domain reflectometry devices. Many of these devices lend themselves to automated electronic data acquisition and recording. As a result, today's geotechnical engineer may have access to a tremendous amount of data relating to a particular facility. Unfortunately, this abundance of data does not immediately lead to increased knowledge – that is, distilled information useful in decision-making. Additional strategies and technologies must be employed to gain the greatest benefit from remote monitoring installations.

## The Need for System Integration

An additional challenge presented by today's array of geotechnical instrumentation techniques is that of system integration. While it is sometimes practical to measure all the desired quantities of a facility using one particular method, it is common for a number of different sensing techniques to be employed in order to obtain a more complete picture of the facility's performance. Data may be collected manually from devices such as piezometers, slope inclinometers, and surveying instruments. These data may be collected by an engineer or technician in the field and recorded with a laptop, handheld device, or paper logbook. A number of different formats exist for recording these data, many of them proprietary. Other data may be collected electronically by an autonomous data acquisition system. Both commercial off-the-shelf and custom-built field computers are used to record data from instruments such as tiltmeters, electronic inclinometers, displacement sensors, and time-domain reflectometry devices. In these cases, the instruments themselves may employ a variety of different analog and digital output schemes, and furthermore, the field computers or data acquisition systems which record the sensor output may employ different data formats, which may be proprietary. A data exchange system must be used in order to glean the most useful knowledge for decision-making from these varied data sources.

The Internet-enabled system for engineering data exchange developed at Northwestern University provides stakeholders with searchable, formatted data from a wide variety of sources on a Web site for easy access.

## THE DATA EXCHANGE FRAMEWORK

The key to effective data exchange lies in providing layers of abstraction between the varied formats in which field data are recorded and the final portable format presented to stakeholders.

Figure 1 shows the Northwestern framework for engineering data exchange. Data are collected in the field either manually or by an autonomous data acquisition system. For large projects, both manually and autonomously acquired data are often available. Once data are gathered, they are sent to a designated computer in a lab or office server space. This polling computer is the gateway through which all data streams enter the electronic data exchange system, and thereby provides a layer of abstraction for the varied incoming data formats.

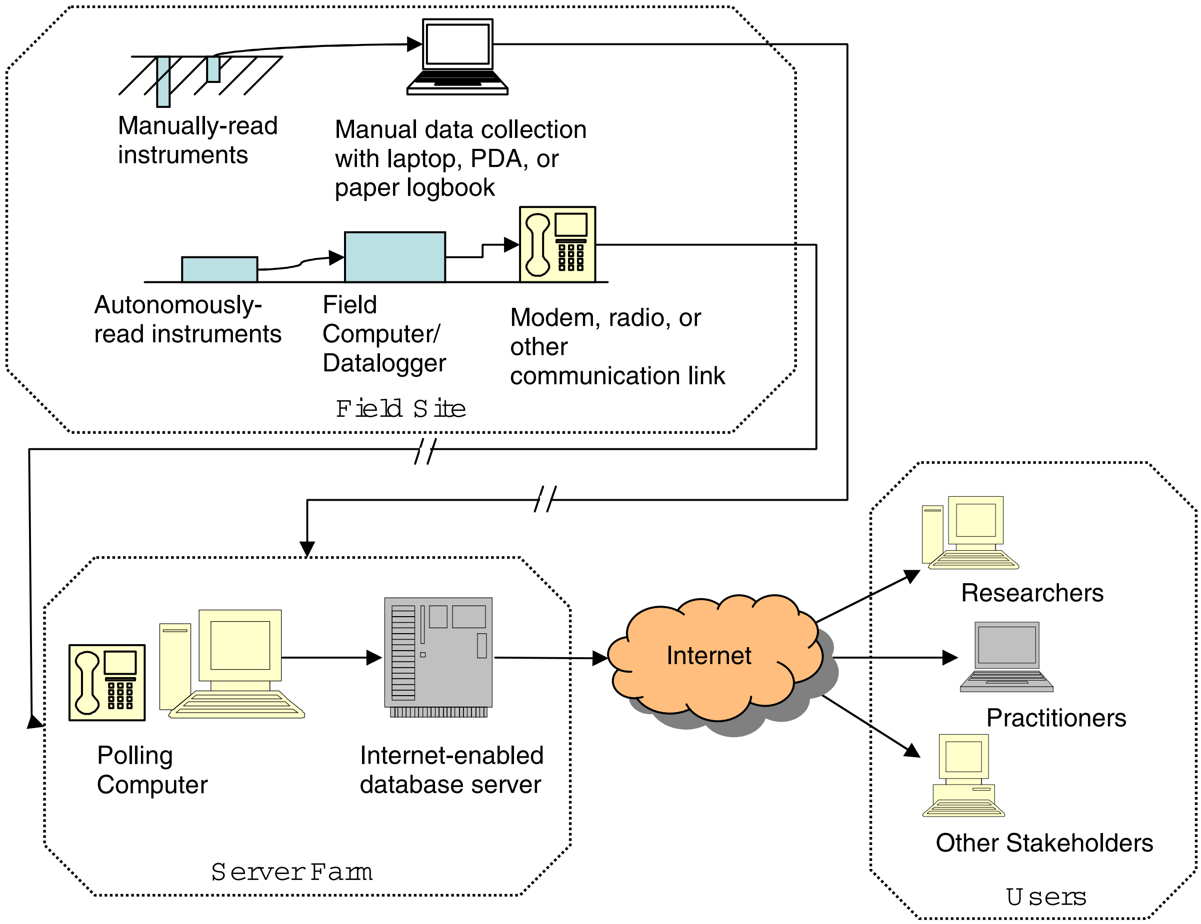


**FIG. 1. Data exchange framework.**

## Polling of Autonomously-Collected Data

Computerized data acquisition systems are most useful when data may be autonomously downloaded on a regular basis by a polling computer in a lab or office. Communication with the field computer may be established by telephone modem, Internet connection, radio/satellite link, or any other digital communication method.

The polling computer has two purposes. First, it establishes contact with field computers by any of the aforementioned methods and downloads data at regularly intervals – typically daily, but sometimes more or less often depending on the nature of the project. Second, the polling computer hosts any proprietary software needed to convert or parse incoming data into a more useful format. Many commercial field computer systems employ a proprietary binary data format that may be read only by the manufacturer's software. In these cases, automation software is used to autonomously invoke the manufacturer's proprietary application, open the new binary data file, and save it in a more universal format, typically tab-delimited text. Some equipment manufacturers have a Web application service that may be used to perform the binary-to-text conversion. In these cases, automation software on the polling computer may be used to autonomously invoke the Web application and save the converted data for further processing.

### Import of Manually-Collected Data

Geomechanical data are often collected from instruments which are manually read using a laptop or handheld computer, an application-specific electronic reader, or pencil and paper. In each of these cases, the data must be converted into a common electronic format for exchange.

The import of data collected via portable computer or application-specific reader is similar to that for remotely-accessible field computers with proprietary data formats; the manufacturer's software is used to export data to a delimited text file which is readily suited for database entry.

Paper logbooks, however, present some difficulty. In some cases, it may be practical to use optical character recognition (OCR) technology to read the logs and generate a digital text file. Where this is not practical, manual data entry must be used. While this is a tedious process, entering the data into a digital file makes it practical to view and share the pencil-and-paper measurements with stakeholders through electronic media.

### Data Exchange and Display via Web Site

The goal of the polling and conversion process is, of course, to import data from the various streams into a relational database through which it might be shared and displayed in readily useful formats. A Web site is an excellent forum for the distribution of engineering data because it provides access to the data set without imposing any particular software or hardware requirements on the user beyond an Internet connection and a Web browser.

### Intermediate Format Abstraction

An intermediate format is employed to provide a layer of abstraction between proprietary data formats associated with various dataloggers and the data archiving and display subsystems. This intermediate format must be readily human- as well as machine-readable. For this reason, intermediate data are written to plain tab-delimited text files such that humans may read the data in columns, while the tabs indicate the end of each field to computer programs used elsewhere in the process. Plain text is not the most efficient way to store numeric data; however, it represents the most universally supported method.

Timekeeping is also an issue in the intermediate format. All timestamps are written as Julian dates representing the local time and the remote site. The Julian date is the number of days that have passed since a particular epoch date; as such, they are easily represented in computing as fixed-point decimal numbers, and are not vulnerable to confusion between varied calendar date formats.

The field-to-Internet data process begins, then, with the conversion of raw data from the dataloggers into this universal intermediate plain-text format. For very small amounts of data, this universal intermediate format might be a sufficient means of storing remote monitoring data. This so-called “flat file” format may be readily imported into the user’s graphing application of choice. However, as data sets grow, flat files quickly become cumbersome and other archiving methods are needed.

**Data Archiving and Display**

Data from the intermediate text format are imported into a relational database program that conforms to the Structured Query Language (SQL) standard. All storage and search functionality is delegated to the relational database. Relational databases are specifically designed for storing and searching large datasets, and the SQL standard provides a clear and widely-supported abstract method of interacting with the database. Data may be retrieved from the database by date range, by event type, or for time periods in which a particular threshold is exceeded on one of the measured channels, for example.

Chadwick et al. (2006) noted an important pitfall in the use of relational databases in report generation: unlike a spreadsheet application, where the plots are directly linked to the spreadsheet data and are automatically updated as data changes, plots generated based on database output must be explicitly regenerated as new data arrives. The framework described in this paper meets that challenge by combining the procedures for data archiving and display into a flexible and scalable Web application. Custom software running on a Web server generates and distributes plots and tables based on data in the relational database. Users receive data as HTML, plain text, and images through their Web browser; no special plug-ins such as Java, JavaScript, Flash, Shockwave, Adobe Reader, or ActiveX are required (Kosnik 2000).

Several Web-based clients for geotechnical information exchange are already available, such as the WebDACS system described by Morris and Farrington (2006). Systems such as WebDACS depend on external plug-ins such as Flash for user input and graphical data display. This dependency can present a number of problems for the user. Such external plug-ins may be unavailable to a stakeholder for a number of reasons: it may be disabled by the IT department, unavailable for a particular operating system or computing platform (handheld computers or PDA’s, for example), or it may be simply too taxing on the user’s available network bandwidth. By strictly adhering to the “lowest common denominator” of Web browser requirements, the Northwestern data display Web site ensures maximum availability to the widest variety of users.

The data search and storage abstraction allows for easy adoption of new field hardware. Internet-enabled remote monitoring installations have been successfully deployed for months and years using both commercial off-the-shelf and custom-built dataloggers (Kosnik 2006).

**Display Modes**

A wide variety of data acquisition modes are supported by the Internet-enabled database. Long-term time histories, where data are collected at regular intervals (hourly or daily, for example) in order to observe trends over weeks, months, or years, are perhaps the most common scheme. Figure FIG. 2 shows a typical long-term time history from a displacement transducer.

It is often helpful to include burst-history data, wherein data are sampled at high frequency whenever a pre-defined threshold is exceeded on one or more sensors. This is especially useful for monitoring or construction and blasting vibrations. The burst waveform may be displayed alone, or the maximum amplitude recorded during burst events may be superimposed on other data, such as long-term time histories. This allows for a visual comparison of long-term and burst data for a given sensor. The Web interface allows the user to search for burst data in a given date range or simply view the latest burst event. A typical search form, event list, and event waveform are presented in Figures 3, 4, and 5, respectively.

Similar search functionality may be used to view data from instruments such as inclinometers and TDR cables. One may either view the latest result, or search by date; it is also possible to view several plots on the same axes such that data may be viewed in historic context, as shown in Figure 6.

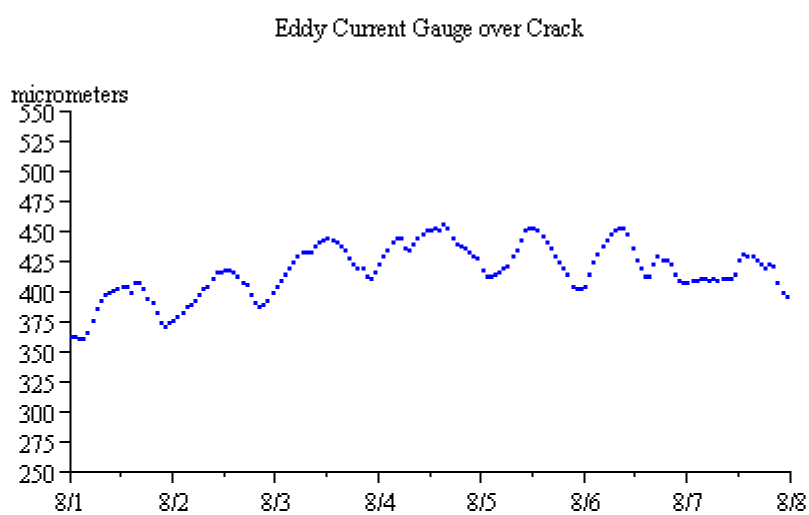


**FIG. 2. Typical long-term time history plot.**


**FIG. 3. Burst event search form.**

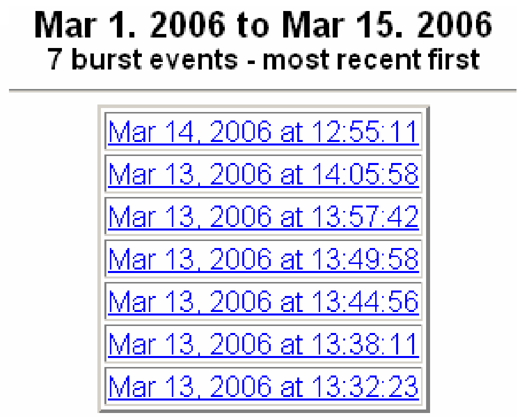


**FIG. 4. Burst event listing.**

## CONCLUSIONS

Modern sensor, communication, and field computer technology make continuous remote monitoring of high-profile or problematic geotechnical facilities a powerful management tool. However, the wide variety of data sources and large volume of data can impede effective distribution and interpretation of field measurements. The techniques for geotechnical data exchange presented here can overcome these challenges by using an Internet-enabled database to archive, plot, and distribute engineering data in near-real time, making it available for management decisions.

While several other techniques have been proposed for Web-based exchange of data from remotely monitored facilities, the Northwestern data exchange framework has been designed from the beginning for maximum scalability in terms of the total volume of data that may be effectively archived and searched as well as maximum usability by clients on a wide variety of operating systems, computing platforms, and Internet connections.

## ACKNOWLEDGEMENTS

The author wishes to thank the faculty and staff of the Northwestern University Infrastructure Technology institute, especially Professors Charles Dowding and Richard Finno, ITI Executive Director David Schulz, and the author's colleagues in the ITI Research Engineering Group: David Prine (retired), Daniel Marron, Mathew Kotowsky, and Daniel Hogan.

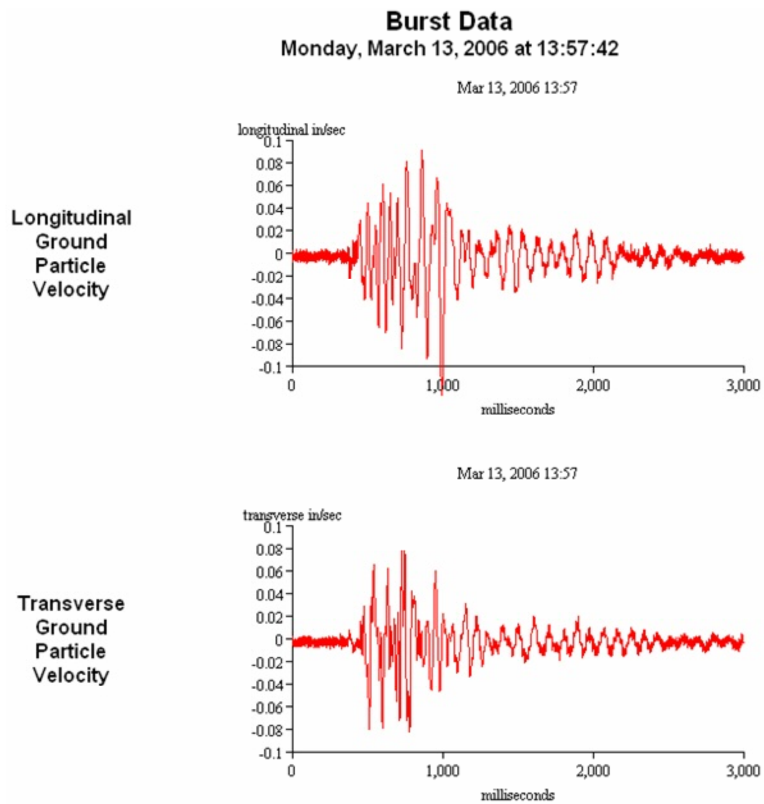


FIG. 5. Burst event waveform.

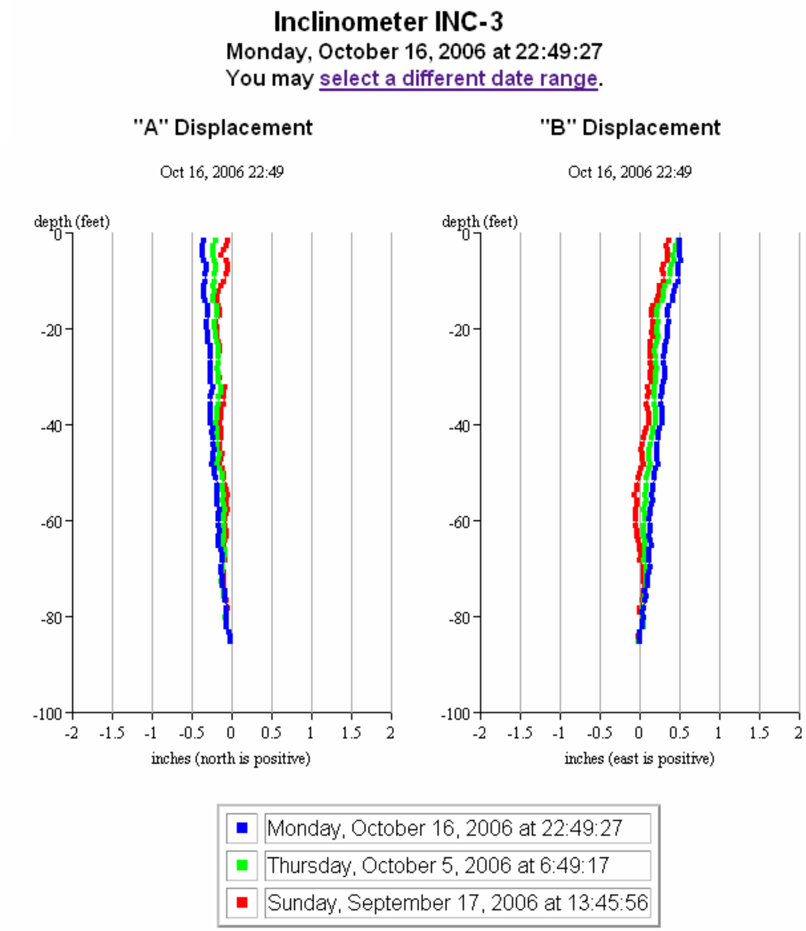


FIG. 6. Inclinometer data.

**REFERENCES**


Chadwick, N., Pickles, C., and Sekulski, E. (2006). "Data Transfer and the Practical Application of Geotechnical Databases." *Proc. Geo Congress 2006*, American Society of Civil Engineers, Reston, Virginia.

Kosnik, D. (2006). "Internet-Enabled Remote Monitoring of Civil Infrastructure with TDR." *Proc. TDR 2006*, Purdue University, West Lafayette, Indiana.

Kosnik, D. (2000). "Server-Side Java Component of Autonomous Crack Comparometer." Infrastructure Technology Institute internal report, Northwestern University, Evanston, Illinois.

Morris, Z. and Farrington, S. (2006). "An IT Architecture for Web-Based Geotechnical Monitoring." *Proc. Geo Congress 2006*, American Society of Civil Engineers, Reston, Virginia.